

学校编码: 10384
学号:

分类号__密级
UDC_____

廈門大學

碩 士 学 位 论 文

基于灰色系统理论的股票市场模型预测及其改进

**Prediction and improvement of stock market model based
on Grey System Theory**

指导教师姓名:

专 业 名 称: 应用数学

论文提交日期: 2017 年 4 月

论文答辩时间: 2017 年 5 月

学位授予日期: 2017 年 6 月

答辩委员会主席: __

评阅人: __

2017 年 4 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘要

随着社会的发展与科学的进步，人们的物质生活水平日益提高，股票投资也逐渐走进千家万户，成为人们生活密不可分的一部分。由于股票投资具有高收益的特点，使得人们对此趋之若鹜，但是伴随着股票的高收益，股票的高风险性也是不可忽视的一点。因此，如何科学地利用股票市场已有的信息来预测股票市场的变化，尽可能地规避风险与提高收益，也成为了一项重要的课题。

本文是基于灰色系统理论和支持向量机理论，以上证指数的周期数据为研究对象来展开研究的。股票市场中作为一种本征能量系统，包含有大量的历史数据，其高度非线性的特点又使得常用的线性回归方法难以奏效，因此我们采用的方法是灰色系统理论中的 GM (1, 1) 模型，为了让模型更加完善，得到合理的结果后对其进行边值修正。在完成预测后，根据已有的等级精度评价标准来给出预测的等级并且判断预测是否合理。为了单一使用灰色预测模型带来的误差偏大的困扰，我们在原有的基础上加入支持向量机的元素，具体的操作表现为在获取残差序列后，利用支持向量回归机对原模型进行残差修正，将两者进行有机结合以期提高预测的精度。在支持向量回归机中，惩罚因子 C 与核函数参数 γ 的变化是影响支持向量回归机效果的重要因素，我们通过对参数进行不断调节选择，尽量保证最好的回归效果。在完成对原 GM (1, 1) 模型的改进步骤之后，将两次预测的结果进行定量地对比分析，发现各项评价指标都得到了明显的提升，从而达到了对原始模型进行改进的目的。

关键词：股市预测 灰色理论 支持向量机

Abstract

With the development of society and the progress of science, people's material standard of living is improving day by day, and the stock investment has gradually entered into thousands of households. Because the stock investment has the characteristic of high income, it makes people rush for it, but with the high income of the stock, the high risk of the stock can not be ignored. Therefore, how to use the existing stock market information to predict the stock market changes, as far as possible to avoid risks and improve returns, has become an important issue.

Based on the grey system theory and the support vector machine theory, this paper studies the periodical data of Shanghai Composite index. As an intrinsic energy system, the stock market includes a large number of historical data, whose characteristics of highly nonlinear make the linear regression method commonly used ineffective, so we consider using boundary value containing the modified GM(1,1) model changes on the stock market for prediction. In the forecast, we are able to give predicted level and determine whether the prediction is reasonable according to the grade the accuracy evaluation of the existing standard. In view of the error of the grey model is sometimes too large, we'd like to use the machine error correction model to improve the prediction accuracy via support vector regression after obtaining the residual sequence. In the support vector regression machine, we can improve the accuracy of the model by adjusting the penalty factor C of the SVR and the kernel function parameter γ . By the improved GM (1,1) forecasting results of the new model, we also carry on the evaluation and comparison with the original model, found that the evaluation indexes are improved, so as to achieve the purpose to improve the original model.

Keywords: Stock market forecast Grey theory Support vector machine

目 录

第一章 绪论.....	1
1.1 股票市场预测的背景及意义.....	1
1.2 股票市场预测的发展概况与研究现状.....	2
1.3 支持向量机简介.....	5
1.4 灰色预测模型简介.....	6
第二章 支持向量机的基础知识.....	7
2.1 经验风险最优与 VC 维.....	7
2.2 结构风险最优.....	10
2.3 SVM 的数学推导.....	13
第三章 灰色系统理论.....	24
3.1 灰色序列生成.....	24
3.2 灰色关联度.....	26
3.3 GM(1, 1) 模型.....	28
第四章 实证分析.....	35
4.1 样本的选取.....	35
4.2 灰色预测模型的评价及其改进.....	35
4.3 模型比较与结论.....	39
第五章 总结与展望.....	40
5.1 工作总结.....	40
5.2 不足与展望.....	40
参考文献.....	42
致谢.....	44

Content

Chapter1 Introduction.....	1
1.1 The Background and Significance of Stock Market Forecast.....	1
1.2 The Development and Research Status of Stock Market	2
1.3 Brief Introduction of Support Vector Machine.....	5
1.4 Brief Introduction of Grey-forecasting Model.....	6
Chapter2 Basic Knowledge of Support Vector Machine.....	7
2.1 Empirical Risk Minimization and VC dimension.....	7
2.2 Structure Risk Minimization.....	10
2.3 Mathematical Derivation of SVM.....	13
Chapter3 Grey System Theory.....	24
3.1 Grey Sequence Generation.....	24
3.2 Grey Relativity Analysis.....	26
3.3 GM(1, 1) Model.....	28
Chapter4 Empirical Analysis.....	35
4.1 Sample Selection.....	35
4.2 Evaluation and Improvement of Grey-forecasting Model.....	35
4.3 Model Comparison and Conclusion.....	39
Chapter5 Summary and Prospect.....	40
5.1 Summary of the Work.....	40
5.2 Deficiency and the Future of Research.....	40
Reference.....	42
Thank.....	44

第一章 绪论

1.1 股票市场预测的背景及意义

社会经济的变化总是让普通大众感到云里雾里，而这些变化通过爆炸式的数据呈现在了我们面前，IT 技术的日益成熟与互联网对万千大众的普及又使得这些数据能够得以不断积累。在这浩瀚如烟海的数据里，我们想做的就是如何从已经积累的数据中，攫取一些重要的信息，让我们能够不至于迷失在数据的海洋中。

随着时代的发展和社会的进步，对股票投资的需求也蒸蒸日上。股票的价格走势直接影响着投资者的经济利益，也影响和反映着国家的宏观经济政策，因而受到人们的广泛关注。股票预测就是参考股票市场数据的历史信息，采用科学可行的预测方法通过对历史数据的分析研究而得到股票数据的未来走势。股票市场体量庞大，因此有存在各式各样的因素影响股票价格的变化，为了方便我们将这些因素分为两大类：基本因素和技术因素。一般地说，基本因素主要包括经济性因素、政治性因素、人为操纵因素和其他因素等，并且这些因素相互交错，特别复杂，最后导致的结果就是它们的“一举一动”都对股价的走势变化有着不可估量的影响，而技术因素则是指影响股票市场价格的各种股票市场操作，让人们耳熟能详的一些名词诸如“追涨杀跌”“买空卖空”都是技术因素对股票市场最根本的影响表现。究其原因不难想到，股民和金融机构是股票市场的基本参与者，他们在股市交易过程中各种买进卖出的行为会积少成多地影响到股价的变化，而股价的变化又反过来影响着人们的交易行为。股市的涨跌变化不仅仅与股市投资者的利益息息相关，同时它也是金融市场中一块巨大的拼图，因此它的变化对金融市场的稳定性。

在传统的数据分析研究领域，我们常常使用包括线性回归、Logistic 回归、多元回归等计量回归方法，而已有的一些研究表明，股市中的大量数据波动是非线性^[2]的状态，如果我们继续使用传统的方法，会导致研究在一开始就出现问题从而得不到真实的结果反馈。因此，我们必须考虑使用一些对非线性数据的研究处理方法。

1.2 股票市场预测的发展概况与研究现状

在金融经济学的发展上,先辈们孜孜不倦地对经济预测进行尝试探索,也为后人开拓发展新的理论提供了重要的经验和思路。在经济预测这颗枝繁叶茂的大树上,股市预测则是它最为粗壮的树枝之一。它的基本思路就是对股票市场中已有的资料信息进行收集处理,然后立足于股市的现状,根据它特有的规律性,运用科学合理的放的方法,对其未来的发展前景进行定性或定量的分析。股市预测是根据过去和现在的股票数值来推出未来的数值,它决定了股票预测的研究对象不是一个特定事件,而是随机的、不确定的事件,这就需要使用与之相匹配的科学预测方法进行合理推断。对于股票预测来讲,其基本特点如下:

(1) 股票市场本质是一个极其复杂的非线性动态系统,虽然近年来对股票市场的研究从未停止,但是并没有一种方法能保证其精确性,所以如何减少预测的误差,保证预测的精度是我们一个重要的任务。

(2) 根据已有的研究,我们知道股票市场具有高度复杂性和非线性的特点,因此对股票市场的预测方法必须是对非线性数据拟合产生良好效果的。但与此同时,对非线性系统理论的研究并不如线性系统理论一般如此完善,许多方法都需要进一步的理论与实践。

(3) 股民作为股票投资者的主体,其具有主观性、分散性、波动性的特点,因此即使对股票市场的历史数据进行了很好的建模,由于股民心理状况、经济状况难以预测,在受到个体对预测结果的影响下,预测结果仍然有很大的可能性令人不甚满意。

(4) 我国作为具有中国特色社会主义的国家,大的经济环境和国家整体的经济实力也对股市的发展有着不可估量的作用,所以我们在做到科学预测的同时,也必须考虑到宏观政策对股票市场的指引以及政治和经济等场外因素的干扰。

考虑到股市预测的复杂性和多变性,我们在预测的时候给出以下三个前提条件:

(1) 有效市场假设:参与市场的投资者有足够的理性,并且能够迅速地对所有市场信息作出合理的反应。

(2) 供求决定假设：供求关系决定交易量和交易价格，并且任何信息都会对供求关系产生影响。

(3) 历史相似原则：未来股市的一切变动趋势都能在过去的历史演变中都能找到规律。

相对于商品经济而言，股票市场的历史就短暂得多，而在这短短一百多年的历史当中，许多让人们眼花缭乱的技术分析方法也随着证券市场运作体系的完善而不断为人们所熟知。。

在 1900 年到 1902 年, 作为《华尔街日报》的创始人和首任编辑查尔斯·道写了许多社论, 讨论股票投机的方法。1902 年查尔斯逝世, 同年 S. A. 纳尔逊将他的评论收集整理并出版《股票投机的基础知识》, 这本书也就是后来鼎鼎大名的作为技术分析基础的道氏理论 (Dow Theory)。19 世纪 20 年代福布斯杂志的编辑理查德·夏巴克, 在道氏理论的基础上, 挖掘出“股价平均指数”对于单个股票的重要信息。1939 年, 美国证券分析专家 R. N. 艾略波特利用道琼斯工业平均指数 (Dow Jones Industrial Average, DIJA) 作为研究工具, 发现股价变化的波动情况并提出了“波浪理论”。而后起源于日本德川幕府时代的 K 线图逐渐走上股市这个大舞台, 由于其具有直观、立体感强、携带信息量大等特点, 逐渐成为技术分析的主流。有了 K 线图之后, 随之而来的趋势分析、移动平均法等技术面分析方法也都像雨后春笋般纷纷涌现。这些分析方法都是立足于图表, 而图表信息的给人的直观感受远远强于一系列单纯的数字, 使得它们在当时的大环境下也取得了一定的成功。但是这种方法缺点也跟它们的优点一样明显, 那就是股票市场的信息过于繁杂, 只是观察图表的变化是人们的一种主观直觉与臆测, 对于其合理性与科学性, 许多人始终抱有怀疑的态度。

到了 20 世纪 60 年代, 人们发现, 经济预测的本质其实就是时间序列预测。时间序列又称为动态数列, 是指将同一统计指标的数值按照发生时间的先后顺序排列而成的数列, 它的一个典型特征就是相邻的数值之间具有一定的依赖性。为了更好的研究这种性质, 各种时间序列模型被人们相继提出, 伴随着的对模型的性质和意义的研究也在不断深入当中。对于金融时间序列的研究基本上分为两种, 一种方法是立足于最基本的经济发展规律, 建立金融时间序列服从的数学模型, 像资本资产定价模型 (CAPM)^[3]、套利定价理论 (APT)^[3]、期权定价模型^[3]等。这些

理论都很成功,但是他们的前提都是建立在种种理想的假设上,而通常市场的实际情况却与这些假设有着很大的出入,所以这些理论在实证研究中并不能让人满意。另一种方法则是抛开经济学与金融学的束缚,回到最原始的数据当中,从数学中概率统计的角度来理性地分析这些数据未来可能的变化。与第一种方法的理论性过强而实证性不足相反,它常常为人所诟病的就是缺乏完整的理论基础,但是种种数据实验表明从统计学的角度来进行分析验证往往能达到更好的效果,也更贴合股价数据运行的轨迹。而且,统计学经过了几百年的发展,各项理论也在不断的完善当中,数理统计学与社会统计学的分歧相比统计学诞生之初也在逐渐缩小当中。更为重要的一点是,在统计方法中有着合理的检验评价体系,这对于评价模型的好坏有着至关重要的作用。

在对时间序列的研究之初,人们把主要的研究精力都投入在自回归移动平均模型^[4] (ARMA) 之中。这是一种结构简单的线性模型,模型之中的统计推断理论也趋于完善,所以得到了广泛的应用。但是该种线性模型的构造条件过于理想,许多现实生活中的产生的数据难以在该模型中反应出来。穷则思变,为了更加贴合实际应用,人们转而开始对非线性时间序列进行大量深入的研究并取得了不斐的成果。在这些耀眼的成果中,最具有划时代意义的就是罗伯特·恩格尔于 1982 年提出的自回归条件异方差模型^[5] (ARCH),他在对大量的数据与模型进行研究后发现预测误差的方差变化会受到诸如宏观政策、政府财政收入等许多场外因素的干扰,并且方差之间的变化存在一定的相关性。自回归条件方差模型的提出就是刻画这种相关性,它不再将方差看作常量而是认为其随时间的变化而变化,这个思路为时间序列模型的构造提供了新的方向,并且成功克服了线性模型在局部数据不稳定的缺点,在实际应用中理论与实际情况也更为相符,应用也更为广泛。

在这之后,混沌理论也开始蓬勃发展并不断完善并且为预测研究打开了一扇新的大门,同时也使得原来许多被人们不可能预测的系统变为可能。混沌理论的优点在于对系统进行量化分析的同时,还带有对系统质性的思考^[8]。在这个大环境下,由中邓聚龙教授在 20 世纪 80 年代首先提出并创立了灰色系统理论,主要用于解决一些包含未知因素的特殊领域问题,在社会、经济、农业、生态等许多系统中都得到了广泛的应用。分形学鼻祖曼德尔布罗特认为股票市场具有一定的自相似性,他在研究了股价的变动规律之后,发现股价变动在各个时期的分布具

有一定的相似性，从而论证了自己的结论。而后科学家们在对股票市场时序的研究中计算出了股票系统的李雅普诺夫指数 (Lyapunov Index)，认为股票市场也存在有混沌现象，因此利用小波理论对金融时间序列进行预测也变成了可能。

这些年来，随着机器学习和神经网络理论发展的日益完善，其在 Pattern Recognition 和 Artificial Intelligence 等方面已经取得了优秀的业绩，并且开始广泛应用在经济、金融等领域。神经网络在非线形拟合方面具有一枝独秀的优势，无论多么复杂的非线性关系，都能被神经网络映射。得益于计算机技术的飞速发展及学习规则的简单，神经网络也容易在计算机上进行实现。尽管它的优点不计其数——很强的鲁棒性、记忆能力、非线性映射能力以及强大的自学习能力，但是在统计小样本数据问题时，神经网络也存在欠学习问题。1995 年，Vapnik 和 Corrina Cortes 提出了支持向量机 (Support Vector Machine, SVM) 的概念，主要的用途是用于分类和回归分析。相比与神经网络，支持向量机具有更好的数学基础，同时泛化能力也更强。支持向量机的稳定性使其能够避免神经网络中遇到的欠学习或过学习的问题，在利用核函数的时候也能够很好的解决“维数灾难”的问题。

1.3 支持向量机简介

20 世纪 60 年代，统计学习理论诞生，几十年的发展使之建立在一套比较坚实的理论基础上。它从统计学的角度证明了有限样本和经验风险和真实风险之间的差异，并首次引入置信区间的概念，创建了结构风险最优理论，为各类机器学习算法提供了统一的评估框架。在 1992-1995 年间，基于统计学习的理论基础发展出了一种新的通用学习方法——支持向量机。可以说支持向量机是统计学习理论在算法领域应用的集大成者。支持向量机的提出，一举解决了第二代神经网络的结构选择和局部最小值（过拟合、欠拟合）等问题，使第二代神经网络进入了又一个低潮期。以统计学习为理论基础的支持向量机现在被应用于机器学习的各个领域^[9]，具体包括文本分类、图像识别、手写数字识别、生物信息学等等。

1.4 灰色系统理论简介

控制论学者 W. R. 艾什比用黑箱 (Black Box) 形容内部信息缺乏的对象和系统。为此，我们用“黑”表示信息缺乏，“白”表示信息完全。信息不充分、不完全称为“灰”。信息不完全的系统，称为灰色系统或者简称灰系统 (Grey System)。

“信息不完全”，一般指

- (1) 系统因素不完全明确；
- (2) 因素关系不完全清楚；
- (3) 系统结构不完全知道；
- (4) 系统的作用原理不完全明了。

信息不完全，是灰色系统的特征。在灰色系统中，我们无法得到完整的信息，也就无法获得所有的数据，那么对于灰色系统理论的研究就是建立在少数据分析的基础之上。少数据中，要求包含现实信息。在我们的认知体系内，即时少量的数据，只要它所蕴含的信息是真实的，那么它一定会表现出不同程度的现实规律。因此，灰色系统理论的目的就在于用刷子逐步刷掉事物表面所覆盖的“灰”，从而使其能恢复本来面目的“白”。现实规律的呈现方式是数值的变化与时间的分布，我们将区间内的数值变化的分布与时间变化的分布定义为灰过程。

广义下的能量系统，基本都带有灰系统本身的特征。我们整个人类社会是能量系统，有了社会属性后人们从简单的以物换物到现在逐渐完善的金融体系所构成的经济系统也是一个大型的能量系统，股票系统作为经济系统的一部分自然也带有能量系统的特征。能量系统的特征是具有一定惯性，我们都知道“质量”越大，惯性也就越大。那么对于这些“质量”或着“能量”大的系统，在发生状态运动的某个邻域内，其将来的运动轨迹必定能从过往的运动状况中变得有迹可循。灰色系统诞生的意义就是利用逻辑完备的数学理论，对系统在自然状态下的惯性变化进行预测。

由于具备了以上的种种特点，灰色系统理论在诸如气象预测、商业预测、工农业产值预测等等领域都有着极其实用的价值，除了预报功能以外，它还有着规划决策的功能，比如计划生产部门如何实施完善可持续发展的生产计划，矿藏勘探部门对开发采集现状进行评估以及开发采集收益进行预判，经济管理部门制定

合理稳定的宏观政策并对宏观经济市场进行调控管理。一言以蔽之，灰色系统理论是一座建造在社会科学与自然科学之间的桥梁，能够将抽象复杂的系统以简单明了的形式表现出来，并且在定性分析的基础上转而进行定量分析，给人客观公正、科学合理的印象，是人类社会各个大型部门与系统用来进行分析预测、决策控制的一种杰出思想理论。

第二章 支持向量机的基础知识

2.1 经验风险最优与 VC 维

由于了总体真实分布的困难（实际上是不可能的），我们只能从总体中抽取若干有代表性的对象，这些对象被称为总体的样本。样本可以理解为总体的若干个观测值，因为是观测值，所以表现形式是静态的、有限的。在数学上可以计算出这些样本的分布，作为事物总体的近似模型或着经验模型。样本来源于总体，那么样本的分布必然与总体的分布有着毋庸置疑的相似性或趋同性，这是我们做样本统计所立足的基本思想。但是，仍然需要衡量经验模型与真是模型之间的误差，为此我们提出一个新的概念，叫做风险。当我们从训练集得到一个分类器之后，这个分类器是以经验模型的分布作为总体分布的，那么这个分类器的误差就称为经验风险^[13]。为了综合评估的方便，我们给这些误差函数起一个新名字：损失函数(Loss Function)。

假设样本： $(x_1, y_1), \dots, (x_n, y_n) \in R^n \times R$ ，对于离散样本的经验函数可以写为：

$$R_{emp}(\alpha) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, \alpha))$$

其中 $R_{emp}(\alpha)$ 就是所谓的经验风险。 $f(x_i, \alpha)$ 是用来最小化风险 $R_{emp}(\alpha)$ 的目标函数， $L(y, f(x, \alpha))$ 就是损失函数，它表示对于每个 x_i ，其标签 y_i 与目标函数 $f(x_i, \alpha)$ 之间的偏差。

我们的目的是为了损失函数尽可能地最小，但是在实际分类训练的过程中，如果只是让损失函数达到数学意义上的最小化，它的分类效果却常常不尽如人意，无法满足全局最优的要求，而且还常常出现过拟合现象^[14]。因此，人们很自然地将目光转向统计方法。人们希望利用统计学找到一种方法来衡量样本所表达的规律与总体规律之间的差距，这也就是所谓学习一致性的问题。于是统计学习理论发展起来了。

统计学习的第一个定义就是要解决经验风险与真实风险的衡量问题。

设 $Q(x, \alpha)$ 是对给定的独立同分布观测样本，使经验风险函数 $R_{emp} = \frac{1}{n} \sum Q(x_i, \alpha)$ 最小化的函数。如果 $R(\alpha_n)$ 与 $R_{emp}(\alpha_n)$ 依概率 P 收敛至同一个极限，即

$$R(\alpha_n) \xrightarrow{n \rightarrow \infty} \inf_{\alpha \in \Lambda} R(\alpha), R_{emp}(\alpha_n) \xrightarrow{n \rightarrow \infty} \inf_{\alpha \in \Lambda} R(\alpha)$$

则经验风险最小原则 (Empirical Risk Minimizing, ERM) 对函数集 $Q(x, \alpha)$ ， $\alpha \in \Lambda$ 概率分布函数 $F(x)$ 是一致的。其中 \inf 表示函数的下确界， $R(\alpha_n)$ 表示真实风险^[15]。

从学习一致性的角度来看，当样本数目趋近于无穷大时（大数定理），最小的经验风险一定能够收敛于最小的真实风险。只有满足这一条件，才能保证在经验风险最小化下取得的最优结果能够代表真实风险的最优结果，如图所示：

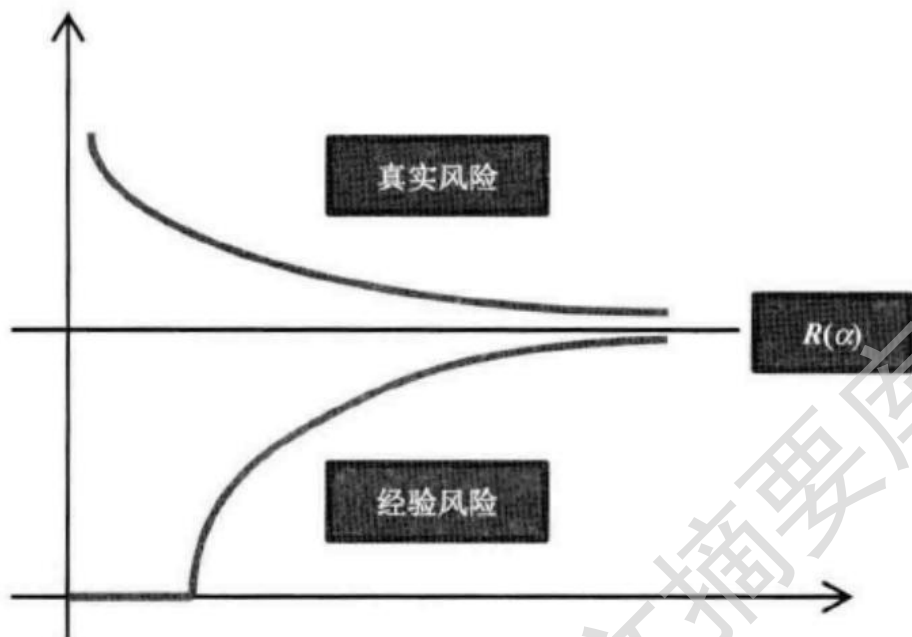


图 2-1 真实风险与经验风险

进一步，Vapnik 还给出了学习理论的关键定理。

设函数集 $Q(z, \alpha), \alpha \in \Lambda$ ，满足条件：

$$A \leq \int Q(z, \alpha) dF(z) \leq B$$

那么 ERM 原则一致性的充分必要条件是： $R_{emp}(\alpha)$ 在如下意义上一致收敛于 $R(\alpha)$ ：

$$\lim_{n \rightarrow \infty} P\{\sup(R(\alpha) - R(\alpha_n)) > \varepsilon\} = 0, \forall \varepsilon > 0$$

其中 P 表示概率，sup 表示函数的上确界。

这个定理被称为学习理论的关键定理，在统计学习理论中具有十分重要的地位。它没有用经验风险去逼近真实风险（因为很困难），而是通过经验风险最小化函数来逼近真实风险最小化函数。这就是关键定理的绝妙之处——我们需要解决的就只是一个收敛问题而不是原来的学习一致性问题。这样，我们在经验风险与真实风险之间构造了一座桥梁，即经验风险最小化原则符合学习过程一致性的条件是， $Q(z, \alpha)$ 函数集中最差的那个函数。

那么如何找到这个最差的函数呢？关键定理没有直接给出方法，但是提供了对 $Q(z, \alpha)$ 的一个度量方法，这就是 VC 维。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库